Journal of Statistical Planning and Inference 179 (2016) 15–21

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Augmented Plackett–Burman designs with replication and improved bias properties

Lu Shen[a], Max D. Morris[b,*]

[a] American Research Institutes, Washington, D.C., United States
[b] Department of Statistics, Iowa State University, United States

ARTICLE INFO

Article history:
Received 15 October 2014
Received in revised form 9 May 2016
Accepted 24 July 2016
Available online 8 August 2016

Keywords:
Foldover
Main effect plans
Orthogonal design
Resolution III
Resolution IV

ABSTRACT

Plackett–Burman designs are among the most popular small, two-level experimental plans used in engineering and industrial applications, due primarily to their relatively small size and orthogonal structure. Two important practical limitations of these designs are (1) the lack of replicate points from which experimental error might be directly assessed, and (2) the potential for serious estimation bias if interactions among the experimental factors are present. We describe a strategy for augmenting these designs that adds replicate runs and reduces potential bias, while preserving first-order orthogonality.

1. Introduction

Since the publication of "The Design of Optimal Multifactorial Experiments" by Plackett and Burman (1946), the Plackett–Burman (P–B) designs have been widely applied; Bakonyi et al. (2011), Vatanara et al. (2007), and Liu and Tang (2010) are three examples of applications in which these designs are prominent. P–B designs are two-level plans that are orthogonal under a first-order (main effects) model, and for any number of experimental factors, contain the smallest number of runs for which this is possible. Plackett and Burman tabulated designs for n runs where n is a multiple of 4 from 8 to 100 except for 92, for which the n-run design can be used for experiments including up to $n-1$ factors. (A design for $n = 92$ was subsequently published by Baumert et al., 1962.) Many popular textbooks on experimental design (e.g. Myers et al., 2009) reproduce the "P–B patterns" – the first row of the design matrix – for the relatively smaller values of n, from which these designs can be completed by a cyclic construction. P–B designs are of Resolution III, i.e. the main effects of individual factors are not aliased with each other, but are aliased with linear combinations of two-factor interaction effects. P–B designs are especially popular in screening experiments because they allow detection of large main effects with small designs, under the assumption that all interactions are negligible.

P–B designs are two-level factorial plans that have two very desirable properties. First, they are orthogonal main-effects designs, and so coefficient estimates for the first-order model have the smallest possible variances for the given design size. Second, among two-level orthogonal main-effects plans, they have the smallest possible number of runs for $f = n-1,\ n-2,\ n-3$, or $n-4$ factors. Hence, P–B designs are especially popular in applications where a main effects model is tentatively assumed to be adequate and relatively precise estimates of the effects are desired, given operational constraints limiting the number of experimental runs. However, in the form in which they were originally presented, P–B designs also

* Correspondence to: Department of Statistics, Iowa State University, Ames, IA 50011, United States. Fax: +1 515 294 4040.
E-mail address: mmorris@iastate.edu (M.D. Morris).

http://dx.doi.org/10.1016/j.jspi.2016.07.003
0378-3758/

have two properties that are less appealing: (1) they contain no replicated experimental runs, and so provide no "pure error" degrees of freedom for estimating the experimental error variance, and (2) if higher-order terms (i.e. interactions) are actually present, these can lead to seriously biased estimates of main effects. Both of these weaknesses are directly related to the especially small size of the P–B designs. In fact, P–B designs are most notable for providing orthogonal main effect estimates, in many cases with a smaller number of runs than would be required by regular orthogonal fractional factorial designs for which n must be a power of 2. It should not be surprising that a design of this size that is optimal in one sense is somewhat less than perfect in others.

In many cases, an estimate of the error variance based on true replicate runs, or better protection against bias associated with higher-order terms, is important enough that an experimenter is willing to increase the size of the experiment beyond the number of runs required by a P–B design. If all the factors in the experiment are continuous, multiple center points can be added to provide an estimate of error variance. However, this strategy does not work if at least one of the factors is not continuous or cannot be set to a third intermediate level, and it is this situation we address here.

If only one of the two undesirable properties (lack of replication, or bias vulnerability) is to be addressed, the design may be doubled in one of two different ways, both of which maintain the orthogonal structure of the design:

1. Each of the runs of the P–B design may be executed twice. This results in "pure error" degrees of freedom equal to the number of runs in the original P–B design, and so provides a solid basis for estimating error variance. However, the bias structure of this design is the same as with the original P–B design, i.e. each main effect estimate is potentially biased to the same extent by the same higher-order terms as with the original P–B plan.
2. The P–B design may be doubled by augmenting with the "foldover" of each run in the original plan, i.e. the run for which each factor originally at the high level is set to the low level, and each factor originally at the low level is set to the high level; Box and Hunter (1961). This process eliminates any biasing of main effects by two-factor interactions not included in the model. However, none of the foldover runs are replicates of any of the original runs in a P–B design, and so this augmented design – like the original – provides no degrees of freedom for a pure error estimate of error variance.

As a compromise approach, we describe a doubling strategy based on P–B designs that (1) provides pairs of replicate runs and hence pure error degrees of freedom for estimating the error variance, and (2) minimizes, to the extent possible, the potential biasing of main effect estimates by two-factor interactions. These designs are constructed by doubling the original P–B (not its foldover), and rearranging the columns in the second half of the design in such a way as to jointly maximize the number of replicate run pairs and minimize a measure of potential estimation bias.

2. Design construction

2.1. $f = n - 1$

We begin our discussion by considering design construction for $f = n - 1$, the largest number of factors that can be accommodated in a P–B design of size n. (We address the construction of designs in $f = n - 2$, $n - 3$, and $n - 4$ factors in Section 2.2.) Using the standard ± 1 coding for two-level designs, let $\mathbf{D}$ be the n-by-f design matrix for the P–B design with the maximal number of factors. We refer to the first row of $\mathbf{D}$ as the "P–B pattern", provided by Plackett and Burman in an extensive table. The next $n - 2$ rows contain all cyclic permutations of this P–B pattern, and the elements of the last row of $\mathbf{D}$ are all -1. So, the first row of $\mathbf{D}$, as tabulated by Plackett and Burman, is sufficient to construct the entire design matrix.

For example, the P–B pattern for $f = 7$ is:

$$+1, +1, +1, -1, +1, -1, -1,$$

and following the cyclic construction pattern described above, the design matrix is:

$$\mathbf{D} = \begin{pmatrix} +1 & +1 & +1 & -1 & +1 & -1 & -1 \\ -1 & +1 & +1 & +1 & -1 & +1 & -1 \\ -1 & -1 & +1 & +1 & +1 & -1 & +1 \\ +1 & -1 & -1 & +1 & +1 & +1 & -1 \\ -1 & +1 & -1 & -1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 & -1 & +1 & +1 \\ +1 & +1 & -1 & +1 & -1 & -1 & +1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{pmatrix}.$$

Let $\mathbf{X_1}$ be the n-by-$(f + 1)$ model matrix for the main effects model including an intercept, i.e.

$$\mathbf{X_1} = (\mathbf{1} \quad \mathbf{D}).$$

Letting $\boldsymbol{y}$ be the n-element vector of responses, the least-squares estimate of the vector of $f + 1$ parameters in the first-order model is:

$$\mathbf{b} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\boldsymbol{y} = \frac{1}{n}\mathbf{X}_1'\boldsymbol{y},$$

the latter form following because the design is orthogonal.